ARMA Models – Prediction and Impulse Response Function

Katalin Varga (Ph.D.)

BUTE

2016 Autumn

Multivariate ARMA Models

Consider a multivariate time series:

$$x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}$$

The AR(1) is $x_t = \varphi x_{t-1} + \epsilon_t$, reinterpreting this yields:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \varphi_{yy} & \varphi_{yz} \\ \varphi_{zy} & \varphi_{zz} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \delta_t \\ \gamma_t \end{bmatrix}$$

AR: $A(L)x_t = \epsilon_t$

ARMA: $A(L)x_t = B(L)\epsilon_t$

Impulse Response Function

Impulse response: the path of x_t if it is kicked by a single unit shock in ε_t.

For an AR(1) model:

$$x_t = \phi x_{t-1} + \varepsilon_t \text{ or } x_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}.$$

$$\begin{array}{cccccc} \varepsilon_t & 0 & 0 & 1 & 0 & 0 \\ x_t & 0 & 0 & 1 & \phi & \phi^2 \end{array}$$

For a MA(∞) model:

$$x_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}.$$

$$\begin{array}{cccccc} \varepsilon_t & 0 & 0 & 1 & 0 & 0 \\ x_t & 0 & 0 & 1 & \theta_1 & \theta_2 \end{array}$$

Impulse Response Function

For vector variables:

In AR form:

$$A\,(L)x_t = \varepsilon_t, A(0) = I, E(\varepsilon_t {\varepsilon_t}') = \Sigma$$

In MA form:

$$x_t = B\,(L)\varepsilon_t, B(0) = I, E(\varepsilon_t {\varepsilon_t}') = \Sigma$$

$$B(L) \;= A(L)^{-1}$$

$$x_t = B\,(L)Q^{-1}Q\varepsilon_t, \quad x_t \quad = C(L)\eta_t$$

$C(L) \;= B(L)Q^{-1}, \quad Q^{-1}Q^{-1\prime} = \Sigma$, Cholesky decomposition.

Sims orthogonalization- Specifying C(0)

Sims orthogonalization: specifying C(0), so let's take every variable's immidiate response to orthogonalized shocks. Sims: let C(0) be lower triangular:

$C(0) = B(0)Q^{-1} = Q^{-1}$, so the Cholesky decomposition will do.

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} C_{0yy} & 0 \\ C_{0zy} & C_{0zz} \end{bmatrix} \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \end{bmatrix} + C_1 \eta_{t-1} + \cdots$$

Blanchard- Quah orthogonalization : specifying C(1), so we are calculating the long term responses. Let C(1) be lower triangular matrix.

$$\Delta x_t = C(L)\eta_t$$

$$C(1) = B(1)Q^{-1}$$

Prediction- Forecasting

For AR(1): $x_t = \varphi x_{t-1} + \epsilon_t$ we have:

$$E_t(x_{t+1}) = E_t(\varphi x_t + \epsilon_{t+1}) = \varphi x_t$$
$$E_t(x_{t+2}) = E_t(\varphi^2 x_t + \varphi\epsilon_{t+1} + \epsilon_{t+2}) = \varphi^2 x_t$$
$$E_t(x_{t+k}) = \cdots = \varphi^k x_t$$

$$var_t(x_{t+1}) = var_t(\varphi x_t + \epsilon_{t+1}) = {\sigma_\epsilon}^2$$
$$var_t(x_{t+2}) = var_t(\varphi^2 x_t + \varphi\epsilon_{t+1} + \epsilon_{t+2}) = (1 + \varphi^2){\sigma_\epsilon}^2$$
$$var_t(x_{t+k}) = \cdots = \left(1 + \varphi^2 + \cdots + \varphi^{2(k-1)}\right){\sigma_\epsilon}^2$$

$$\lim E_t(x_{t+k}) = \varphi^k x_t = 0$$
$$\lim var_t(x_{t+k}) = \left(1 + \varphi^2 + \cdots + \varphi^{2(k-1)}\right){\sigma_\epsilon}^2 = \frac{1}{1-\varphi^2}{\sigma_\epsilon}^2 = \mathrm{var}(x_t)$$

For MA(1): $x_t = \epsilon_t + \theta\epsilon_{t-1}$ we have:

Homework

Variance Decomposition

In the orthogonalized system we can compute: what percent of the k step ahead forecast error variance is due to which variable.

$$x_t = C(L)\eta_t, \qquad \mathrm{E}(\eta_t \eta_t{}') = \mathrm{I}$$

The one step ahead forecast error variance is:

$$\varepsilon_{t+1} = x_{t+1} - E_t(x_{t+1}) = C_0 \eta_{t+1}$$

$$= \begin{bmatrix} c_{yy,0} & c_{yz,0} \\ c_{zy,0} & c_{zz,0} \end{bmatrix} \begin{bmatrix} \eta_{y,t+1} \\ \eta_{z,t+1} \end{bmatrix}$$

$$C(L) = C_0 + C_1 L + C_2 L^2 + \ldots$$

$$var(y_{t+1}) = c_{yy,0}{}^2 \sigma^2(\eta_y) + c_{yz,0}{}^2 \sigma^2(\eta_z)$$

$$= c_{yy,0}{}^2 + c_{yz,0}{}^2.$$

Granger Causality

„Cause preceeds effect”:

If a random/ unexpected event A forecasts (preceeds) B, then we know, that A „causes” B.

Definition: w_t Granger causes y_t if w_t helps to forecast y_t, given the past y_t:

Consider a vector autoregression:

$$
\begin{aligned}
y_t &= a(L)y_{t-1} + b(L)w_{t-1} + \delta_t \\
w_t &= c(L)y_{t-1} + d(L)w_{t-1} + \nu_t
\end{aligned}
$$

Our definition means that if w_t **does not** Granger cause y_t, if $b(L) = 0$, i.e. if the VAR is in the form:

$$
\begin{aligned}
y_t &= a(L)y_{t-1} + \qquad\qquad + \delta_t \\
w_t &= c(L)y_{t-1} + d(L)w_{t-1} + \nu_t
\end{aligned}
$$

Be careful: Granger causality is not equal with causality!

VARs in State Space Form

Let's rewrite the VAR model in AR(1) form:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \varepsilon_t$$

The so-called companion form of the VAR:

$$\begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \dots \\ I & 0 & \dots \\ \vdots & \dots & \dots \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \\ \vdots \end{bmatrix} + \begin{bmatrix} Q^{-1} \\ 0 \\ \vdots \end{bmatrix} [\eta_t]$$

$$x_t = A x_{t-1} + C\eta_t, \qquad E(\eta_t \eta_t') = \mathrm{I}$$

VARs in State Space Form

The impulse response function: $C, AC, A^2C, \ldots$

$$IR_0 = C, IR_j = AIR_{j-1}$$

If Q^{-1} is a lower triangular matrix, then the first shock influences only the first variable:

$$var_t(x_{t+j}) = \sum_{j=0}^{k-1} A^j CC' A'^j$$

The variance decomposition of the τ-th shock can be calculated by a recursion:

$$v_{i,\tau} = CI_\tau C', \qquad v_{k,t} = Av_{k-1,t}A'$$

Regression

Regression: approximating a random variable Y in $\mathbb{R}^p$ with a function of X, which is in $\mathbb{R}^q$:

$$E(|\boldsymbol{y} - f(\boldsymbol{x})|^2) \to min$$

It is a well-known fact that, the solution of the minimization excercise is the conditional expectation of Y conditioned on X:

$$f(\boldsymbol{x}) = E(\boldsymbol{y}|\boldsymbol{x})$$

It is also well-known from the L^2 theory of random variables that the conditional expectation is a linear function of X if the variables are normally distributed.

Maximum Likelihood (MLE) and Least Square Estimators (LSE)

The model of parametric regression:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{\varepsilon}$$

Where the observations are organized as row vectors of $\boldsymbol{Y}$ és $\boldsymbol{X}$.

$\boldsymbol{B}$ is the matrix of unknown regression coefficients, and $\boldsymbol{\varepsilon}$ is random vector of error terms.

The least square estimator (LSE) of $\boldsymbol{B}$, which is in case of normally distributed random vectors a maximum likelihood estimator (MLE) as well:

$$\widehat{\boldsymbol{B}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

If the inverse of $\boldsymbol{X}^T\boldsymbol{X}$ does not exist, we use the so-called pseudo inverse matrix.

Literature

1. John H. Cochrane: Time Series for Macroeconomics and Finance
2. J. D. Hamilton: Time Series Analysis
3. J. Durbin, S. J. Koopman: Time Series Analysis by State Space Methods
4. Ruey S. Tsay: Analysis of Financial Time Series
5. M. J. Miranda, P. L. Fackler: Applied Computational Economics and Finance
6. Y. Ait-Sahalia, J. Jacod: High-Frequency Financial Econometrics
7. Massimiliano Marcellino: webpage
8. Eric Ghysels: Matlab Toolbox for Midas Regressions
9. Christopher Sims: webpage
10. G. Koop, D. Korobilis: webpage